# No Risk No Trust: Investigating Perceived Risk in Highly Automated Driving

**Mengyao Li**
Georgia Institute of Technology
Atlanta, GA, USA
mli9687@gatech.edu

**Brittany E. Holthausen**
Georgia Institute of Technology
Atlanta, GA, USA
brittany.noah@gatech.edu

**Rachel E. Stuck**
Georgia Institute of Technology
Atlanta, GA, USA
stuck.rachel@gatech.edu

**Bruce N. Walker**
Georgia Institute of Technology
Atlanta, GA, USA
bruce.walker@gatech.edu

## ABSTRACT
When evaluating drivers' trust in automated systems, perceived risk is an inevitable, yet underestimated component, especially during initial interaction. We designed two experimental studies focusing on how people assess risk in different driving environments and how introductory information about automation reliability influences trust and risk perception. First, we designed nine driving scenarios to determine which factors influence Perceived Situational Risk (PSR) and Perceived Relational Risk (PRR). Results showed that participants identified levels of risk based on traffic type and vehicles' abnormal behaviors. We then evaluated how introductory information and situational risk influence trust and PRR. Results showed that participants reported the highest level of trust, perceived automation reliability, and the lowest level of PRR when presented with information about a highly reliable system, and when driving in a low-risk situation. These results highlight the importance of incorporating perceived risk and introductory information to support the trust calibration in automated vehicles.

### Author Keywords
trust in automation, perceived risk, introductory information, simulator study, car-driver interaction.

### ACM Classification Keywords
• **Human-centered computing~User studies** • **Human-centered computing~Empirical studies in HCI** • Human-centered computing~HCI theory, concepts and models • **Applied computing~Psychology**

## INTRODUCTION
Automated vehicles have become widely available to customers with features such as adaptive cruise control and automated lane keeping. As more automated systems are incorporated into new vehicles, automated driving becomes more complex and unfamiliar for the general public. Especially during the first few interactions, it is unfeasible for drivers to understand the entire system[11]. Therefore, a certain level of uncertainty about automation functionality is inherent when interacting with automated vehicles [24]. Trust is a way that humans mitigate uncertainty, so trusting an automated system is essential in uncertain conditions [25]. Thus, to understand trust in car-driver interaction, it is crucial to unravel uncertainty and understand perceived risk during the initial interaction. However, little experimental research addresses how to capture this uncertainty in automated driving context and how risk further influences trust. To address these gaps in the literature, we first established an experimental baseline of various perceived risks in the automated driving environment. Second, we determined how the introductory information provided to drivers, interacting with situational risks, influences perceived risk and trust during the drive.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
*AutomotiveUI '19,* September 21–25, 2019, Utrecht, Netherlands
© 2019 Association for Computing Machinery.
ACM ISBN 978-1-4503-6884-1/19/09 $15.00
https://doi.org/10.1145/3342197.3344525

## BACKGROUND
Trust in automation is a critical factor that determines acceptability, attitude, behavioral intention, and actual use of automated vehicles [8]. Because of this, trust can have substantial impacts on safety and driving performance [8]. Placing an appropriate amount of trust in the machine is crucial during car-driver interaction [18]. An inappropriate amount of trust in the automation can lead to misuse or disuse of automation associated with overtrust and distrust [17]. Only when trust is calibrated to automation capabilities will the functions of automation be maximized and unexpected situations of automation be minimized [13].

A key component of trust in automation is the aspect of 'vulnerability' or 'uncertainty' [15, 16]. Vulnerability and uncertainty are defined as "the trustor must be exposed in some ways" (p.742) [5]. Operating a vehicle has inherent risks, and when a person allows an automated system to perform some or all of the driving tasks, the person becomes vulnerable to any performance problems of the system. Thus, to trust the automated systems, drivers need to willingly take risks when they cannot control the system [15, 16]. Usually, the higher the perceived risk of the automated vehicle, the lower the trust [21, 22, 27]. When a higher level of risk is involved, the greater the trust is needed from the drivers to rely on the automation [16, 23]. Without the potential

outcomes brought by risky situations, drivers could say they have 'confidence' rather than trust in the automated systems [28]. Therefore, it is essential to highlight the presence of risk whenever evaluating trust and how risk influences trust in automated driving.

Trust is dynamic with a high degree of volatility [13]. The development of trust is non-linear with steep growth in the beginning few interactions [2]. Subsequently, the level of trust will become calibrated and stabilize to the level of actual automation functionality if users have an appropriate initial understanding of the system [2, 3]. If drivers are not aware of the limitations of the system before the first use, the trust will be detrimentally decreased when facing system failures [3]. Understanding the initial interaction (i.e., introductory information) becomes critical to foster an appropriate trust and enhance the car-driver interaction. Future automated vehicles are designed for the general public who usually do not have a broad knowledge of the functionality of automation. Thus, the information drivers receive before interacting with automated vehicles is essential to shaping expectations of automation and to further enhance driving safety and performance [14]. It is expected that risk has a greater impact on trust in first interactions, because a deeper level of trust that is less influenced by situational factors is only developed over time [5, 25]. Thus, it is essential to understand how introductory information impacts trust, as well as how drivers' perceived risk during the initial interaction impacts trust.

There are various ways of presenting this introductory information, which may impact the driver's mental model and expectations about an automation's reliability [1, 3, 11]. These have been shown to impact trust and reliance on automation [3, 13]. Introducing sufficient and accurate information could enhance the effectiveness of communication between drivers and automated systems [22]. Prior research has found different results about introductory information in automated driving. Beggiao, Pereira, Petzoldt, and Krems found that if the information was provided about potential limitations of automated systems, it decreased levels of initial trust [2]. If drivers are informed about automation failures beforehand, trust will not decline even after experiencing the failure [2]. Results are conflicting, though, as other research has found that the introductory information did not find significantly impact trust after more exposure to some take-over drives [11]. The effect of initial introductory information on trust was substituted by the consequent driving experience [11]. While trust requires the risk component, no study considers the role and impact of uncertainty feelings when studying introductory information in car-driver interaction. Therefore, we aimed to investigate the effect of introductory information and perceived risk on trust with multiple drives and a period reinforcement of system capability in between drives.

To rigorously assess the impacts of information and perceived risk on trust, we must first establish and evaluate a risk manipulation carefully. Especially in a research context, with ethical considerations, the perceived risk may not be equivalent to that perceived in a real-world driving environment. We should design driving scenarios and ensure that participants perceive various risks in the driving simulator. However, no research has categorized and manipulated risk by levels in a systematic manner in car-driver interaction [25]. Therefore, we designed Study 1 to classify levels of risky situations within specific categories, and investigate factors affecting risk perception in car-driver interaction.

Risk perceived by the driver is an evaluation of both the *probability* of potential negative accidents and the *seriousness* of the potential accidents [16]. For example, driving at high speed decreases the *probability* of avoiding an accident; the *seriousness* of crashing into a truck is greater than crashing into a bicycle. Various driving conditions (e.g., speed, traffic types, abnormal vehicle behaviors) will be designed to cover levels of risk in car-driver interaction.

We evaluated two types of perceived risk in Study 1: perceived *situational* risk (PSR) and perceived *relational* risk (PRR) [25]. PSR is "the belief … that a specific task or context has potential negative outcomes, apart from personal history, or historical knowledge of a human (or automation) in that situation" [25]. In the car-driver interaction, PSR refers to drivers' attitude towards the driving task or context based on the evaluation of potential accidents' probability and seriousness. On the other hand, PRR is "an individual's belief … that interacting with automation with which a user has history has potential negative outcomes" [25]. In the car-driver interaction, PRR refers to drivers' attitude towards the automated vehicle based on both previous and ongoing relationship with automated vehicles. For examples, prior information about automation reliability and previous driving performance will affect drivers' PRR towards automated vehicles. As each of these types has unique relationships with trust [25], it is essential to understand perceptions of perceived relational and situational risk.

The goal of this project was to understand how risk and introductory information influence trust in automated driving. We first investigated various risky situations in the drives, to understand what factors influence two types of perceived risk. Then, we investigated the influences of prior information about automation reliability and situational risk on drivers' trust, perceived relationship risk, and perceived automation reliability.

## STUDY 1: PERCEIVED RISK EVALUATION

In Study 1, there were nine driving scenarios that incorporated different threats in the driving environment; this was to investigate both types of perceived risk. Participants rated their perceived relational risk and perceived situational risk after driving each scenario. We aimed to identify various levels of risks in the driving simulator and what factors impact participants' ratings.

## Method

*Participants* Fourteen undergraduate students (5 females, 9 males) participated in the study. Participants were required to have a driver's license, two years of driving experience, and normal or corrected-to-normal vision. Participants were recruited to investigate factors that can affect drivers' attitude on the automated driving. Participants had licenses for 4.39 years, on average (*SD* = 2.50), and drove an average of 6.86 hours per week (*SD* = 7.75). Participants were 21.31 years of age on average (*SD* = 2.63).

*Apparatus* Drives were simulated using the National Advanced Driving Simulator MiniSim v2.2.1., equipped with three primary displays (approximately 160° view), an instrument panel display, infotainment display, 2.1 audio system, tactile transducer, steering wheel, gas pedal and brake, and adjustable driver's seat (see Figure 1). This simulator is equipped with adaptive cruise control (ACC) and automated lane keeping (ALK) systems. When both of these are activated, the simulated vehicle is at Level 3 automation, as defined by the Society of Automotive Engineers (SAE) [23]. The ALK system controls lateral position of the vehicle; ACC controls longitudinal position of the vehicle. Participants could disable the systems by turning the steering wheel, pressing the gas, or pressing the brake.



**Figure 1. Driving simulator setup.**

| Category | Scenario | Driving Speed | Traffic | Abnormal Behaviors |
|---|---|---|---|---|
| Driving Speed | High Speed | 55 miles/hour | Level C* LOS** in an urban setting. | No abnormal car behaviors. |
| | Medium Speed | 40 miles/hour | | |
| | Low Speed | 25 miles/hour | | |
| Traffic | Trucks | 40 miles/hours | Level C LOS with more trucks and a preceding truck. | No abnormal car behaviors |
| | Cyclists | | Level C LOS with more cyclists and a preceding cyclist. | |
| | Heavy Traffic | | Level E*** LOS with a preceding vehicle. | |
| Abnormal Behaviors | Other Swerves | 40 miles/hours | Level C LOS in an urban setting. | The adjacent vehicle swerves and crosses into the driving lane. |
| | Other Merges | | | The vehicle merges in the front of the driving vehicle leaving a close headway distance. |
| | Subject Swerves | | | The vehicle swerves within the adjacent lane. |

* Level C LOS represents a stable flow or near free flow. The range of traffic density is between 18 to 26.
** LOS: Level of service. This is used to measure the traffic flow. The units of measurement for traffic density are passenger cars per mile per lane (pc/mi/li).
*** Level E LOS represents an unstable flow. The range of traffic density is between 35 to 45.

**Table 1. Descriptions of Driving Scenarios**

*Materials* Three *categories* of driving scenarios were designed by manipulating: driving speed, traffic (volume and type), and abnormal behaviors (see Table 1). In each category, there were three different driving *scenarios*. Thus, participants experienced nine driving scenarios in total. Each scenario lasted for approximately one minute.

*Measures* To assess various risks within a given scenario, two types of perceived risk were measured: perceived situational risk and perceived relational risk. PSR was measured with a one-item scale: "The situation was risky." PRR was measured with a three-item scale: "The device is risky"; "Using the device increases the risk of collision with the preceding vehicle"; and "Using the device forces me to take a lot of risks" [21]. Participants responded to the PSR and PRR items using a continuous slider ranging from zero (strongly disagree) to ten (strongly agree) in order to reflect the continuous nature of the PRR and PSR variables.

### *Procedure*

After participants signed the consent form, they were screened for simulator sickness [7]. Participants completed a demographics and driving experience survey. Following the surveys, participants drove each of the nine driving scenarios, rating their PRR and PSR after each scenario. The *Medium Speed* scenario was always the first drive experienced by

| | | PSR | | | | PRR | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Scenario | *M* (*SD*) | Mean Rank | Chi-Square | Asymp. Sig. | *M* (*SD*) | Mean Rank | Chi-Square | Asymp. Sig. |
| Driving Speed | High Speed | 2.45 (2.04) | 2.42 | 4.00 | .135 | 2.30 (2.37) | 2.33 | 2.51 | .285 |
| | Medium Speed | 1.29 (1.41) | 1.92 | | | 1.36 (1.63) | 1.79 | | |
| | Low Speed | 2.49 (2.97) | 1.67 | | | 2.30 (2.73) | 1.88 | | |
| Traffic | Trucks | 1.56 (1.77) | 1.42 | 10.94 | .004 | 1.78 (2.09) | 1.62 | 12.50 | .002 |
| | Cyclists | 3.85 (2.13) | 2.69 | | | 3.73 (2.20) | 2.77 | | |
| | Heavy Traffic | 2.88 (2.79) | 1.88 | | | 2.23 (2.86) | 1.62 | | |
| Abnormal Behaviors | Other Swerves | 5.74 (2.45) | 2.29 | 8.71 | .013 | 4.48 (1.29) | 2.07 | 19.00 | .000 |
| | Other Merges | 3.66 (2.16) | 1.36 | | | 2.66 (1.96) | 1.14 | | |
| | Subject Swerves | 5.50 (2.19) | 2.36 | | | 5.96 (2.75) | 2.79 | | |

**Table 2. Mean, standard deviation, and Friedman test for all drives on perceived situational risk (PSR) and perceived relational risk (PRR).**

participants. This drive was kept as the first drive for each participant to ensure that their first experience with the automated vehicle was neutral and did not impact their risk ratings on the other scenarios. The remaining eight drives were randomly ordered to avoid order effects.

### Results

Three categories (i.e. Driving speed, Traffic, and Abnormal Behaviors) were tested to determine whether there were significant differences within each category. Higher scores indicate a higher level of perceived risk. For this study, the Cronbach's alpha for the three PRR items was 0.91. Due to a small sample set, the non-parametric Friedman test of differences for repeated measures was used for the analyses (see Table 2).

For *Driving Speed*, there were no significant differences among the three scenarios for both the PRR and PSR scales. There were significant differences between *Traffic* scenarios for both PRR and PSR ratings. For both PRR and PSR ratings, the *Cyclists* drive was rated as the riskiest within the category. For *Abnormal Behaviors*, the differences between three driving scenarios on PSR and PRR ratings are both significant. The ranked levels of PSR and PRR both in the order of *Subject Swerves*, *Other Swerves*, and *Other Merges*. The Friedman tests revealed that participants rated PSR for the scenario *Cyclists* and *Subject Swerve* significantly higher than other scenarios within each category.

The power analysis was conducted in G*Power. For three groups being measured across nine observations, an alpha of .05 and a power of .80, running a repeated measure between subject ANOVA, the required sample size to detect a significant effect is 18 with the actual power of .89.

### Study 1 Discussion

To determine whether participants perceive risk in an experimental setting, and what factors were perceived as risky in Level 3 driving situations, we designed nine scenarios to manipulate and evaluate two types of risks: perceived situational risk and perceived relational risk. Participants were able to distinguish and identify high versus low levels of perceived risk associated with the driving scenarios within certain categories. For example, the drives *Cyclists* and *Subject Swerve* were rated significantly higher than other scenarios within each category.

We also found out that drivers showed various responses to different characteristics of automated driving. Driving speed was surprisingly not a significant factor affecting drivers' perceived risk in automated driving. Most participants complained about the low driving speed throughout the experiment. Due to the low fidelity of the driving simulator, it is likely that participants have a mismatched expectation of speed between driving in a simulated driving environment and driving in a car-racing video game.

For *Traffic*, drivers reported that using automated vehicles on the road with a preceding cyclist was highly risky both situationally and relationally. In the *Cyclists* scenario, the seriousness of a crash with the preceding cyclists was much higher than with a car because cyclists are more likely to get injured in an accident. If the automated vehicle failed to adapt, the total harm of an accident when adding the harm towards the cyclist would be much higher.

For *Abnormal Behaviors*, drivers considered the longitudinal event, *Other Merges* scenario perceived as less risky when compared to the lateral event, as experienced in *Subject Swerves* and *Other Swerves*. It is possible that the experience of vehicles suddenly cutting in happens more often in daily life compared vehicles swerving.

This is the first study to look at perceived situational and relational risk. Further exploration is needed to understand how PSR and PRR relate to trust and each other. A highly reliable automated system or a driver's trust in the system will mitigate the PRR level even with a high level of situational risk on the road. On the other hand, if a driver distrusts the system, the PRR level will be high no matter whether situational risk is present. Further research could also focus on other situational risks to be incorporated into the scenarios (e.g., weather, highway vs. city). A systematic evaluation of scenarios to quantify situational risks in the driving environment is necessary.

This study provided information on participants ratings of different levels of risks and major factors influencing their perceptions. Although there are limitations such as complex differences between the simulation and real-life driving and a small sample size, this study provided the first successful manipulation of perceived risk in the automated driving environment. Study 1 ensured the presence of the risk component, which built the foundation for Study 2 to further investigate drivers' trust in car-driver interaction.

## STUDY 2: EFFECT OF INTRODUCTORY INFORMATION & SITUATIONAL RISK ON TRUST

Since effects of risk on trust are more sensitive in the first few interactions, the goal of Study 2 was to investigate the effects of prior information about automation reliability on drivers' attitudes towards the automated vehicles in two levels of situational risk (i.e., high vs. low). Study 2 used a 2 x 2 mixed design with the same driving simulator using Level 3 automation as Study 1.

We manipulated introductory information on automation reliability. *Reliability* is defined as the consistency of automation performance in situations that involve some degree of dependence and risk [4]. By giving drivers information on automation reliability before driving, drivers will form a mental measure of automation's future consistency of performance [11, 13, 15]. Two levels of information on automation reliability were used, high and low, requiring a between-subjects design. The high-reliability group of participants was informed that "the automated system was highly reliable, which means the system would perform well throughout the drives." The low-reliability group was told that "the automated system reliability level was low, which meant the system would break down 25% of the time". The actual automation reliability was the same among the two groups.

Building on the results from Study 1, we also manipulated two levels of situational risk (i.e., high and low) associated with drives as a within-subject variable. Each participant drove three drives. The drive *Medium Speed* is the baseline control drive. A baseline drive was adopted in order to show the anchoring effects from the initial introduction. We used two drives from the *Abnormal Behavior* category: *Other Merges* (low) and *Other Swerves* (high). The scenario, *Subject Swerves*, was not employed because the vehicle's reliability level was low in this drive, which would influence the effects of introductory information on automation reliability. Thus, risk levels in drives were strictly controlled by manipulating the situational risk around the car.

Dependent variables were: self-reported trust level; perceived automation reliability (PAR); and perceived relationship risk (PRR).

With this study design, we investigated the influences of introductory information on automation reliability interacting with the situational risk level of drives on the reported trust, perceived automation reliability, and perceived relationship risk. We postulated the following hypotheses:

**H1:** High-reliability information would lead to a higher level of trust, PAR, and lower level of PRR.

**H2:** Higher situational risk level of drives would lead to a lower level of trust, PAR, and a higher level of PRR.

**H3:** There would be an interaction effect of introductory information and situational risk. Specifically, high-reliability information mixed with a low-risk drive would lead to the highest level of trust, PAR, and the lowest level of PRR.

### Measures

#### *Trust*

Trust was evaluated using a 12-item Trust in Automation Scale [10], which participants responded to with a continuous slider ranging from 1 to 7 [10]. A higher total score indicated a higher level of trust.

#### *Perceived Automation Reliability*

Reliability was evaluated using a 4-item scale which is adapted from a subscale of Reliability/Competence of Trust in Automation (TiA) [12]. The range of the scale is -3 (strongly disagree) to +3 (strongly agree).

#### *Perceived Relational Risk*

The perceived relational risk was assessed with the same scale as used in Study 1.

### Procedures

After each participant signed the consent form, they completed the Simulator Sickness Screening protocol. Then participants provided demographics and their driving history. Next, the experimenter read the introductory information on automation reliability, which based on the randomly assigned group (high or low). Participants then completed trust, PRR, and PAR measures as a manipulation check. Participants were given instructions about how to use the automation in the driving simulator and asked to complete the baseline drive (i.e., Medium Speed). Two different levels of situational risk were randomly ordered after the baseline drive. After each drive, participants completed the same set of surveys measuring trust, PRR, and PAR. After the study, the experimenter debriefed the participant. For details see Figure 3.

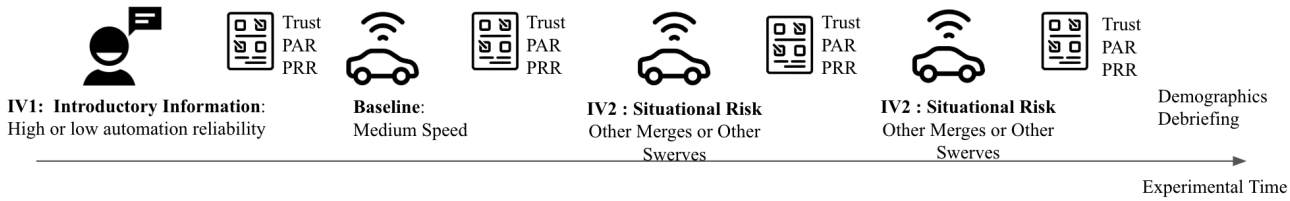


**Figure 3. Experiment Setup.**

| Dependent Variables | α | 95% CI | *df* | Group | *F* | *p* | $\eta_p^2$ |
|---|---|---|---|---|---|---|---|
| Perceived Automation Reliability | 0.89 | 0.20 - 0.96 | (1,61) | Situational Risk | 18.855 | .001 | .236 |
| | | | | Introduction | 19.706 | .001 | .244 |
| | | | | Situational Risk ×Introduction | 4.834 | .032 | .073 |
| Trust | 0.90 | 4.43 - 5.05 | (1,61) | Situational Risk | 19.639 | .001 | .244 |
| | | | | Introduction | 25.131 | .001 | .292 |
| | | | | Situational Risk ×Introduction | .148 | .702 | .002 |
| Perceived Relationship Risk | 0.88 | 3.56 - 4.65 | (1,61) | Situational Risk | 35.571 | .001 | .368 |
| | | | | Introduction | 15.157 | .001 | .199 |
| | | | | Situational Risk ×Introduction | .179 | .673 | .003 |

**Table 4. Two by two ANOVA result with Cronbach's alphas values of dependent variables.**

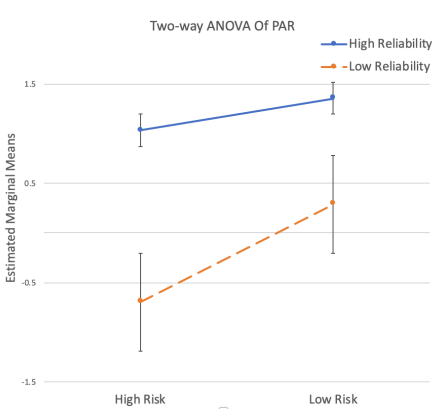


**Figure 4. Two-way ANOVA of drivers' PAR**




**Figure 5. Two-way ANOVA of drivers' trust**

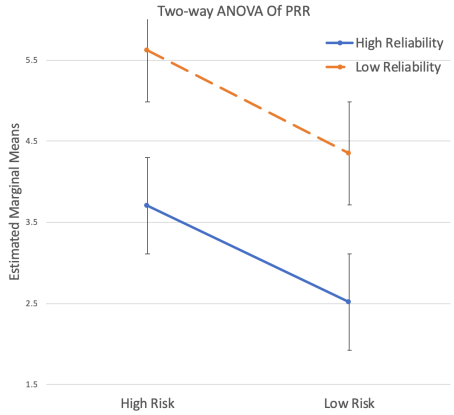


**Figure 6. Two-way ANOVA of drivers' PRR**

## Participants

Seventy undergraduate students were recruited based on the same pre-screening requirements as listed for the first study. Seven participants were eliminated because of uncompleted data due to technical failures during the study. Participants included 27 females and 36 males between the ages of 18 to 28. The average age of the sample was 19.44 (*SD* = 1.36). The average number of years holding a license was 3.14 (*SD* = 1.35), with the participants driving an average of 5.68 hours per week (*SD* = 4.16).

## Results

A 2x2 ANOVA was used to test the effects of automation reliability described in the introduction and situational risk level of drives. All effects were statistically significant at the .05 level. Dependent variables were perceived automation reliability, trust, and perceived relational risk.

*PAR.* Figure 4 depicts the values of perceived automation reliability in high/low reliability group experiencing high/low risk level of drives. The high/low reliability represents the reliability described in the introduction. The high/low risk represents the situational risks presented in two drives. There was a significant interaction between the effects of reliability information and risk level of the drives ($F_{1,61} = 4.834$, $p < .05$). A two-way analysis of variance yielded a main effect for the between-subject variable reliability ($F_{1,61} = 19.706$, $p < .001$), such that the average PAR score was significantly higher for high reliability group ($M = 2.39$, $SD = 1.42$) than for low reliability group ($M = -.20$, $SD = 1.35$). A main effect between the within-subject variable situational risk was found ($F_{1,61} = 18.855$, $p < .001$). indicating that the PAR was significantly higher for low risk level drive ($M = .81$, $SD = 1.46$) than for high level drive (*M*

= .15, *SD* = 1.63) (See Table 4). No ceiling or flooring effects were found for PAR measurements.

*Trust.* Figure 5 depicts the values of trust in high/low reliability group experiencing high/low risk level of drives. A two-way analysis of variance yielded a main effect for the between-subject variable reliability ($F_{1,61} = 25.131$, $p < .001$), such that the average trust level was significantly higher for the high reliability group (*M* = 5.41, *SD* = 1.19) than for the low reliability group (*M* = 4.14, *SD* = 1.04). A main effect between the within-subject variable situational risk ($F_{1,61} = 19.639$, $p < .001$), indicating that the trust was significantly higher for low risk level drive (*M* = 4.99, *SD* = 1.28) than for high level drive (*M* = 4.48, *SD* = 1.28). However, the interaction between effects of reliability information and risk level of drives was not significant (See Table 4).

*PRR.* Figure 6 depicts the values of perceived relational risk in high/low reliability group experiencing high/low risk level of drives. A two-way analysis of variance yielded a main effect for the between-subject variable reliability ($F_{1,61} = 15.157$, $p < .001$), such that the average PRR score was significantly higher for the low reliability group (*M* = 4.98, SD = 3.13) than for the high reliability group (*M* = 3.11, *SD* = 2.55). A main effect between the within-subject variable situational risk was found ($F_{1,61} = 35.571$, $p < .001$), indicating that the PRR was significantly higher for the high-risk level drive (*M* = 4.72, *SD* = 2.35) than for the low-level drive (*M* = 3.49, *SD* = 2.20), which also confirms the results from Study 1. However, the interaction between the effects of reliability information and risk level of drives was not significant (See Table 4).

### Study 2 Discussion

Study 2 investigated the effects of introductory information on automation reliability and situational risk levels of drives on trust, perceived relational risk, and perceived automation reliability. Hypothesis 1 predicted drivers' attitude towards the automation could be manipulated by prior information of automation reliability. Hypothesis 2 tested whether the situational risk in drives could impact drivers' attitudes. Results showed that both introductory information on automation reliability and situational risk level of drives had significant effects on drivers' trust, perceived relational risk, and perceived automation reliability. We also investigated whether the situational risk levels and introductory information would interact and influence drivers' attitudes. This interaction effect was significant for perceived automation reliability (see Figure 4), but not for trust and perceived relational risk (see Figure 5, 6).

We can consider the reasons for the significant interaction effect for perceived automation reliability but not trust and perceived relational risk. First, as an underlying dimension of trust in automation [12], PAR demonstrated a clearer connection with the independent factor, the prior information of automation reliability, whereas for trust and PRR are moderated by other factors such as individual differences in propensity to trust. Second, the situational risk in drives could be viewed as the automated system is not highly reliable and adaptive to handle all critical events, which causes a direct decrease in PAR. We highlighted that the uncertainty of automation (in this study, uncertainty in the introductory information and situational risk drives) only impacts perceived automation reliability rather than trust in automation [25]. The interaction effect was only significant on PAR rather than trust, confirming the statement in the introduction that human-automation confidence differs from the human-automation trust. Thus, for future studies, it is necessary to carefully measure and distinguish between the concepts of perceived automation reliability and trust. One key component is risk or uncertainty involved, after evaluating the possible outcomes brought by risky situations, drivers are actually trust the automated system rather than being confident about the system reliability [28].

These results confirm the effects of preliminary reliability of automated systems [3, 6, 11]. When drivers were told about low reliability, they had lower trust for the automated system. In contrast, in Körber, Baseler, and Bengler [11] introduction materials consisted of three parts (a trust promoted/lowered video, text, and drive) and were presented only once in the beginning of the study. A significant difference for trust was not found. Our findings show that it is effective to change drivers' mental model of automation reliability through a quantitative way (i.e., the system will break down 25% of the time) and with a periodical manner (i.e., repeat every time before each drive). This study suggests that manuals for new interfaces periodically presenting quantitative information might be helpful for drivers to form an appropriate mental model of automated systems [29].

This study also showed that the prior information about the automation and situational risk in drives was closely related to self-reported PRR levels, which confirming results from Study 1 about the impact of scenario design, and its relationship with PRR.

#### *Limitations and Future Directions*

One constraint is the difference between the simulation and real-life driving. To the extent possible, future research should explore introductory information and perceived risk in real-driving scenarios (especially perception of risk as the consequences are real). Also, future research should include a broader population to improve the generalizability of the results. Especially in older populations, as age is an important factor affecting risky driving behavior, and perception of risk between the older driver and younger drivers [9]. Third, the differences between PSR and PRR have not been extensively studied in the literature since they have just recently been identified. Our future research is working to build validated scales of both constructs.

### Conclusion

This study provides guidelines on how to manipulate and incorporate risk into experimental studies in the automated driving context. Importantly, we highlighted the importance of evaluating various types and levels of risks as components

of trust. Without taking risk into account, the validity of trust cannot be guaranteed. In most cases, perceived automation reliability is measured, rather than trust in automation. Presenting drivers with appropriate information will significantly influence their perceived automation reliability and trust in automation. Therefore, integrating perceived risk, especially in the early stages of car-driver interaction, will be extremely important to ensure a better understanding of drivers' trust in future automated systems.

## REFERENCES


[1] Bailey, N.R. and Scerbo, M.W. 2007. Automation-induced complacency for monitoring highly reliable systems: The role of task complexity, system experience, and operator trust. *Theoretical Issues in Ergonomics Science*. (2007). DOI:https://doi.org/10.1080/14639220500535301.

[2] Beggiato, M. et al. 2015. Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour*. 35, (2015), 75–84. DOI:https://doi.org/10.1016/j.trf.2015.10.005.

[3] Beggiato, M. and Krems, J.F. 2013. The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F: Traffic Psychology and Behaviour*. 18, (2013), 47–57. DOI:https://doi.org/10.1016/j.trf.2012.12.006.

[4] Choi, J.K. and Ji, Y.G. 2015. Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *International Journal of Human-Computer Interaction*. (2015). DOI:https://doi.org/10.1080/10447318.2015.1070549.

[5] Corritore, C.L. et al. 2003. On-line trust: Concepts, evolving themes, a model. *International Journal of Human Computer Studies*. (2003). DOI:https://doi.org/10.1016/S1071-5819(03)00041-7.

[6] Forster, Y. et al. 2018. Calibration of Trust Expectancies in Conditionally Automated Driving by Brand, Reliability Information and Introductionary Videos. *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '18* (New York, New York, USA, 2018), 118–128.

[7] Gable, T.M. and Walker, B.N. 2013. Georgia Tech Simulator Sickness Screening Protocol. (Oct. 2013).

[8] Ghazizadeh, M. et al. 2012. Extending the Technology Acceptance Model to assess automation. *Cognition, Technology and Work*. 14, 1 (2012), 39–49. DOI:https://doi.org/10.1007/s10111-011-0194-3.

[9] Gold, C. et al. 2016. Taking Over Control From Highly Automated Vehicles in Complex Traffic Situations. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 58, 4 (Jun. 2016), 642–652. DOI:https://doi.org/10.1177/0018720816634226.

[10] Jian, J.-Y. et al. 2004. Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics*. (2004). DOI:https://doi.org/10.1207/s15327566ijce0401_04.

[11] Körber, M. et al. 2018. Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics*. 66, (2018), 18–31. DOI:https://doi.org/10.1016/j.apergo.2017.07.006.

[12] Körber, M. 2018. Theoretical considerations and development of a questionnaire to measure trust in automation. *In Proceedings 20th Triennial Congress of the IEA. Springer.* (2018), 13–30.

[13] Lee, J.D. and See, K.A. 2004. Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 46, 1 (2004), 50–80. DOI:https://doi.org/10.1518/hfes.46.1.50_30392.

[14] Mayer, R.C. et al. 1995. *An Integrative Model of Organizational Trust*.

[15] Merritt, S.M. and Ilgen, D.R. 2008. Not All Trust Is Created Equal: Dispositional and History-Based Trust in Human-Automation Interactions. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 50, 2 (2008), 194–210. DOI:https://doi.org/10.1518/001872008X288574.

[16] Numan, J.H. 1998. Knowledge-based systems as companions: Trust, human computer interaction and complex systems. [S.n.].

[17] Parasuraman, R. 1997. Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*. 39, 2 (1997), 230–253. DOI:https://doi.org/39(2), 230-253.

[18] Parasuraman, R. and Riley, V. 2006. Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. (2006). DOI:https://doi.org/10.1518/001872097778543886.

[19] Pederson, B.K.M.K. et al. 2018. Simulations and Self-Driving Cars: A Study of Trust and Consequences. *HRI '18 Companion* (2018).

[20] Perkins, L. et al. 2012. Designing for Human-

Centered Systems: Situational Risk as a Factor of Trust in Automation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. (2012). DOI:https://doi.org/10.1177/154193121005402502.

[21] Rajaonah, B. et al. 2008. The role of intervening variables in driver-ACC cooperation. *International Journal of Human Computer Studies*. (2008). DOI:https://doi.org/10.1016/j.ijhcs.2007.09.002.

[22] Reeves, B. and Nass, C. 1996. The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places (CSLI Lecture Notes S). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. (1996). DOI:https://doi.org/10.1109/MSPEC.1997.576013.

[23] SAE 2013. Summary of Levels of Driving Automation for On-Road Vehicles. *Center for Internet and Society, Stanford Law School*. (2013). DOI:https://doi.org/10.1021/nl072974w.

[24] Satterfield, K. et al. 2017. The Influence of Risky Conditions in Trust in Autonomous Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 61, 1 (Sep. 2017), 324–328. DOI:https://doi.org/10.1177/1541931213601562.

[25] Stuck, R.E., & Walker, B.N. The Importance of Incorporating Risk into the Study of Human-Automation Trust. *Human Factors*.